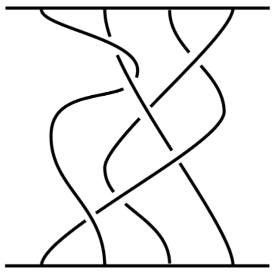
KTH Engineering Sciences

The braid group, representations and non-abelian anyons



OSKAR WEINBERGER

Bachelor's thesis

SA104X Degree Project in Engineering Physics, First Cycle
Department of Mathematics
KTH, Royal Institute of Technology
Supervisor: Douglas Lundholm

May 2015

Abstract

This bachelor's thesis concerns unitary linear representations of the braid group, motivated by their connection to two dimensional particle statistics and certain quasiparticles called non-abelian anyons. Particle statistics in two dimensions is related to the braid group via the fundamental group of the configuration space of indistinguishable particles in two dimensions, and non-abelian anyons correspond to non-commutative unitary representations of the braid group.

In the aim of understanding these connections and studying the mathematical possibilities for non-abelian anyons, algebraic and topological definitions as well as results about braids and their representations are presented and investigated. The focus is on representations of low dimension, and a characterisation of low-dimensional irreducible complex representations is analysed. The unitarisability of such representations and the consequences for non-abelian anyons are then considered.

Contents

Chapter 1

Introduction

The usual way to describe the statistics of identical particles in quantum mechanics is to consider permutations of the particle indices. As the particles are identical, any physical observable should remain unchanged under such a permutation. This does not necessarily require the wave function to be unchanged under the exchange, as physical observables do not depend on the absolute phase of the wave function. In introductory textbooks on quantum mechanics, it is therefore postulated that a system of identical particles may either have a symmetric or an antisymmetric wave function. This is also in agreement with what is observed in nature, namely bosons and fermions, and the postulate seems justified.

However, multiplication with any phase, not just 1 and -1, would meet the physical requirement of unchanged observables, and in this setting there seems to be no convincing argument as to why only symmetric and antisymmetric wave functions would be possible. The idea of particle exchange as merely a permutation of indices also appears rather contrived, as the particles are identical to begin with and the permutation of indices therefore corresponds to nothing physical.

Indeed, it is possible to adopt a different formalism where the exchange of two particles does have a physical meaning. This was first done by Leinaas and Myrheim in [13], where the exchange is described as an adiabatic, continuous transformation, with the particles changing place via continuous paths. They also required that paths differing only slightly, i.e. being continuously deformable to one another, should yield the same result. Continuity corresponds then in part to the physical requirement that the particles are distinct and may not coincide.

In this formalism, they could explain rather than postulate that wave functions must indeed be either symmetric or antisymmetric under particle exchange. In other words, particles are either bosons or fermions, and are said to obey Bose-Einstein and Fermi-Dirac statistics respectively. However, they also noted that the argument was valid only in three dimensions and higher, and that in two dimensions arbitrary phases are possible. Particles obeying such intermediate statistics are now called *anyons*, a name first coined by Wilczek in [25], referring to their property of having any phase. Such particles are necessarily quasi-particles, i.e. emergent properties of a system, as fundamental particles are intrinsically three-dimensional. For a review of the physics of anyons, see [18].

Although two-dimensional physical systems can not in principle occur, a system may still be subjected to a potential effectively restricting it to two dimensions. For a potential $V(x, y, z) = U(x, y) + \phi(z)$ the Hamiltonian splits after separation of variables, and the eigenfunctions factor into a one-dimensional and a two-dimensional part. If the energy level separation of the z-component is sufficiently large, the system will essentially be two-dimensional. It is in this sense that two-dimensional systems can exist.

The essential difference between particle exchange in two and three dimensions may be seen in the simple example of a two particle system. Place the origin at one of the particles, so that an exchange is represented by the other particle traversing around the origin to minus its initial position. Exchanging twice is then represented by a loop around the origin. In our example, the distinctness of the particles corresponds to removing the origin. In three dimensions, the space is still simply connected, and the loop may be continuously deformed to its starting point without the particles coinciding. Hence after exchanging twice, the wave function should obtain its initial value. In two dimensions however, this is not possible, and no number of exchanges generate a loop deformable to a point.

A distinction is traditionally made between abelian and non-abelian anyons. Abelian anyons are particles with intermediate statistics where exchanges of particles commute, i.e. the order of exchange has no effect on the wave function. For non-abelian anyons however, the total phase change of the wave function under particle exchange may depend on the order of exchange. The difference is due to the dimension of the wave function. Anyons with a one-dimensional wave function will be abelian, while anyons with higher-dimensional wave functions will in general be non-abelian.

As a simplified example of how multicomponent wave functions might arise, consider an electron trapped in the plane. Assume also that the plane is penetrated by a discrete set of magnetic flux tubes, or laser beams. The single electron Hamiltonian will then depend on the positions of the tubes or beams in the plane, and we consider these to be parameters. In general, the Hamiltonian of the electron may depend on a set of parameters $x_1, ..., x_k$ of positions in the plane.

Assume further that the ground state of this Hamiltonian is degenerate with degree n. The ground state vector ψ may then be written $\sum_{i=1}^{n} c_i\psi_{0,i}$ for a basis $\{\psi_{0,i}\}$ of the degenerate eigenspace. The basis $\{\psi_{0,i}\}$ and the c_i:s will in general depend on the parameters, and changing $x_1, ..., x_k$ will yield a different basis $\{\psi'_{0,i}\}$ and set of coefficients c'_i. Thus as a function of $x_1, ..., x_k$, the ground state wave function takes values in $\mathbb{C}^n$. Note that since the basis varies, the c_i:s and the c'_i:s can not in general be compared. However, if the parameters are changed and then returned to the initial configuration, the basis may be chosen to be the same, and the coefficients be compared. Typically, the magnetic flux tubes are considered as (quasi)particles of the system, and it is exchange of these particles that may exhibit anyonic statistics. An example of this may be found in for example [4].

Although there exists a possibility for different statistics in two dimensions, the question arises whether physical systems exist that realise anyonic statistics. Indeed, abelian anyons are widely believed to have been experimentally observed and have been conjectured to play a central role in explaining some cases of the so called fractional quantum Hall effect (FQHE), see [12]. There is also evidence suggesting that non-abelian anyons too exist in systems displaying FQHE. Furthermore, systems of non-abelian anyons are serious candidates for building stable quantum computers, so called topological quantum computers, see [11], [21] or [24].

There is also the question how a many-particle system of anyons behave. In the bosonic and fermionic cases, there is Bose-Einstein and Fermi-Dirac statistics, and there the existence of an exclusion principle in the latter case is central. As anyons interpolate between these two statistics, the question arises whether there exists some sort of exclusion principle for anyons. Work regarding this has mostly focused on the abelian case, and abelian anyons have been shown to exhibit localised exclusion, see for example [15], [16] and [17]. Although not a focus of this thesis, an initial question was whether there exist such exclusion principles also in the non-abelian case. In order to answer such questions, the so called representations corresponding to non-abelian anyons must first be examined.

The mathematical treatment of particle statistics goes via the configuration space and its fundamental group, capturing the notions discussed above of a system of distinct particles and deformation of loops respectively. The effect on the wave function under particle exchange is then given by the linear representation of the fundamental group, and the adiabatic nature of the exchange requires this representation to be unitary. In two dimensions, the relevant group is the braid group, a generalisation of the symmetric group, and the study of anyons and their statistical properties is hence directly linked to representations of the braid group.

The aim of this thesis has been to investigate the mathematical possibilities for non-abelian anyons, and thus to study representations of the braid group related to wave functions with more than one component. Relevant definitions and results on representations and the braid group have been collected, as well as the necessary tools from from linear algebra, topology and abstract algebra. Typically, physicists study representations of dimension growing exponentially with the number of particles, see [21]. However, it is natural to ask what the simplest non-abelian unitary representations are for a given number of particles. The main focus has therefore been on low dimensional representations, especially the so called Burau representation, and the unitarisation of such representations.

Chapter 2

Preliminaries

For an excellent introduction to abstract algebra, see [20]. For basic definitions on groups, rings and modules, and a standard reference in the subject, see [6]. A few additional definitions are collected below.

Definition 2.1. *The hermitian transpose of a complex matrix M, denoted M^*, is the conjugate transpose of M, i.e. $M^* = \bar{M}^T$.*

Definition 2.2. *A matrix U is called unitary if $UU^* = U^*U = I$.*

Definition 2.3. *A matrix H is called hermitian if $H^* = H$.*

Definition 2.4. *The minimal polynomial m of a matrix M is the monic polynomial p of least degree such that $p(M) = 0$.*

Least degree is meant in the sense that any polynomial p such that $p(M) = 0$ has degree greater than or equal to the degree of m. As the name suggests, the minimal polynomial is unique. Furthermore, it has the following important properties:

(i) m divides any polynomial p such that $p(M) = 0$.
(ii) M is diagonalisable if and only if $m(M)$ factors into distinct linear factors.

Definition 2.5. *The direct sum of a family $\{M_i\}_{i \in I}$ of modules is the set $\bigoplus_{i \in I} M_i$ of sequences (m_i) where $m_i \in M_i$ and only finitely many of the m_i:s are nonzero.*

The direct sum $\bigoplus_{i \in I} M_i$ is a subset of the direct product $\prod_{i \in I} M_i$. In the case that the indexing set I is finite, the two sets are the same. A sum $M_1 + ... + M_k$ of modules is called direct if there exists a bijection $\bigoplus_{i=1}^k M_i \to M_1 + ... + M_k$, i.e. every element in $M_1 + ... + M_k$ can be written uniquely as a sum $\sum m_i$ with $m_i \in M_i$.

Definition 2.6. *A $\mathbb{C}$-algebra is a unital ring $\mathcal{A}$ together with a ring homomorphism $f\colon \mathbb{C} \to \mathcal{A}$ mapping $1 \mapsto 1_{\mathcal{A}}$ such that $f(\mathbb{C})$ is in the center of $\mathcal{A}$.*

An algebra is the extension of a ring with scalar multiplication of some ring. Often the ring is a field, which is taken here to be $\mathbb{C}$. A prototypical example is the algebra of $n \times n$ complex matrices endowed with matrix addition, matrix multiplication and scalar multiplication. Note that every algebra is also a ring, as well as a module. An algebra over a field, such as a $\mathbb{C}$-algebra, is hence also a vector space.

Definition 2.7. *Given a group G, the complex group algebra $\mathbb{C}G$ is the set of formal sums of the form $\sum_{g \in G} c_g g$ with $c_g \in \mathbb{C}$.*

The group algebra may be thought of as polynomials with the indeterminates replaced by group elements.

For the definition of a topological space and basic notions such as continuity, compactness and Hausdorff spaces, see [1]. For the convenience of the reader, the following definitions are given.

Definition 2.8. *A homeomorphism between topological spaces is a continuous bijection with a continuous inverse.*

As such, homeomorphisms are the isomorphisms of topological spaces. Two topological spaces are called homeomorphic if there exists a homeomorphism between them.

Definition 2.9. *An embedding of a topological space X in a topological space Y is a function $f \colon X \to Y$ whose restriction $\tilde{f} \colon X \to f(X)$ to the image of f is a homeomorphism.*

The unit interval $[0, 1]$ is traditionally denoted I.

Definition 2.10. *A topological space X is called a topological interval if it is homeomorphic to I.*

Chapter 3

Braids

3.1 The Artin braid group

Definition 3.1. *The braid group B_n is the group defined with generators $\{\sigma_1, \sigma_2, ..., \sigma_{n-1}\}$ and relations*

$$\sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1} \quad 1 \leq i \leq n-2$$
$$\sigma_i \sigma_j = \sigma_j \sigma_i \quad |i-j| \geq 2$$

The braid group B_n is called the braid group on n strings, and is usually pictured using so called braid diagrams. In the diagrams, the generators σ_i are twists of adjacent strings, braiding the $(i+1)$th string above the ith. The inverse of each generator is simply to braid in the other direction. An example is given in Figure 3.3. In Figures 3.1 and 3.2, the generators and their relations are shown.

It should be noted that there are several conventions regarding braid diagrams. In the convention used here, a braid is considered to go from top to button. In physics, where the braids are thought of as pictures in space time with time going up, the converse is most often used. Furthermore, the generators may be taken instead as the inverses of those used here, with σ_i being the ith string over the $(i+1)$th. Regardless of such choices, the group depicted is B_n as defined above.

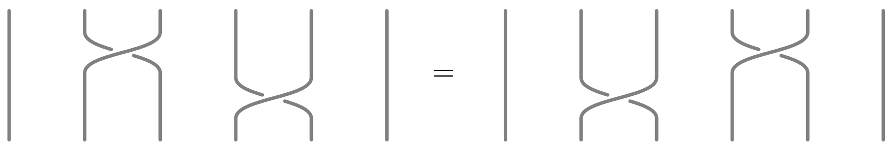


Figure 3.1. Braid diagram showing $\sigma_i \sigma_{i+2} = \sigma_{i+2} \sigma_i$

There is a natural connection between the symmetric group and the braid group; a braid element on n strings induces a permutation of the n starting points. Mapping every braid to its associated permutation generates a surjective homomorphism from B_n to S_n. The kernel of this homomorphism is called the pure braid group, the set of braids with every string starting and ending at the same index.

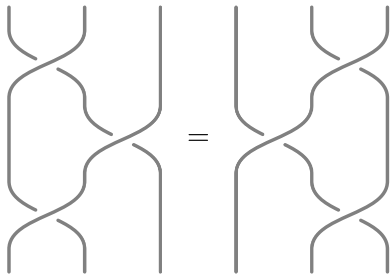


Figure 3.2. The second relation $\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}$

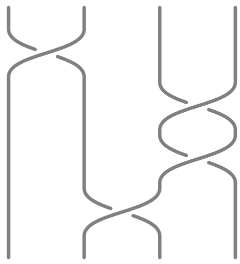


Figure 3.3. A braid diagram of $\sigma_1\sigma_3\sigma_3\sigma_2$

Definition 3.2. *The pure braid group, denoted P_n, is the kernel of the surjective homomorphism p given by*

$$\begin{array}{rccc} p\colon & B_n & \to & S_n \\ & \sigma_i & \mapsto & s_i \end{array}$$

for $i = 1, ..., n-1$, where the s_i:s are the standard generators of S_n.

The symmetric group has a standard presentation with the generators s_i very similar to the standard presentation of the braid group, namely

$$\begin{array}{rr} s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1} & 1 \leq i \leq n-2 \\ s_i s_j = s_j s_i & |i-j| \geq 2 \\ s_i^2 = 1 & \forall\ i = 1, ..., n-1. \end{array}$$

Algebraically then, the symmetric group differ from the braid group only in the additional relation $s_i^2 = 1$.

The presentation given here is known as the Artin presentation, given first by E. Artin in [2]. It defines the braid group in purely algebraic terms, and in this setting, the braid group is referred to as the Artin braid group. As we will see, other presentations with fewer generators are possible. Furthermore, various geometric and topological constructions of groups turn out to be isomorphic to the Artin braid group, offering pictures and visual interpretations of the braid group. For a standard reference in the theory of braids, see [3].

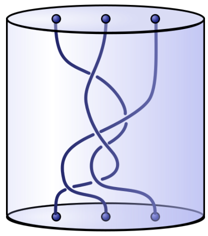


Figure 3.4. A geometric braid on 3 strings

3.2 Geometric braids

Already the braid diagrams in the previous section dealt with a geometric interpretation of the braid group. This section makes this notion explicit.

Definition 3.3. *A geometric braid on n strings $b \subset \mathbb{R}^2 \times I$ is a union of n disjoint topological intervals such that*

(*i*) *each string is mapped homeomorphically onto I by the projection* $\mathbb{R}^2 \times I \to I$

(*ii*)
$$b \cap \left(\mathbb{R}^2 \times \{0\}\right) = \{(1,0),(2,0),...,(n,0)\} \times \{0\}$$
$$b \cap \left(\mathbb{R}^2 \times \{1\}\right) = \{(1,0),(2,0),...,(n,0)\} \times \{1\}$$

The first requirement is roughly that each string crosses every plane $\mathbb{R}^2 \times \{t\}$, $t \in I$ in exactly one point. The second requirement means that the geometric braid induces a mapping of the points $\{(1,0),(2,0),...,(n,0)\} \times \{0\}$ to the points $\{(1,0),(2,0),...,(n,0)\} \times \{1\}$, and hence a permutation of the indexes $\{1,2,...,n\}$. The choice of these sets of points as starting and ending points are of course arbitrary, and could be replaced with any two discrete subsets of the real plane of n distinct points.

To give these geometric braids a group structure, we first need the following definition.

Definition 3.4. *Two braids a and b are called isotopic, or isotopy equivalent, if $\exists F \colon a \times I \to \mathbb{R}^2 \times I$ such that*

(*i*) F *is continuous*

(*ii*) $F_s \colon a \to \mathbb{R}^2 \times I,\ x \mapsto F(x,s)$*, is an embedding whose image is a geometric braid* $\forall s \in I$

(*iii*) $F(a,0) = a,\ F(a,1) = b$

The relation of isotopy between geometric braids defines an equivalence relation, and we can consider the set of equivalence classes of braids, called braids on n strings. This set is then made into a group by defining multiplication of two braids by concatenation. This is stated more explicitly in the next proposition.

Proposition 3.1. *The braids on n strings becomes a group under the multiplication of geometric braids given by*

$$b_1 b_2 = \{(x, y, t) \in \mathbb{R}^2 \times I\} \text{ where } (x, y, 2t) \in b_1 \text{ if } 0 \leq t \leq 1/2 \text{ and } (x, y, 2t-1) \in b_2 \text{ if } 1/2 \leq t \leq 1.$$

The group of braids on n strings is denoted $\mathcal{B}_n$.

When drawing geometric braids, one usually considers their associated braid diagrams as in section 3.1, i.e. their projections onto $\mathbb{R} \times \{0\} \times I$. Here, if two strings cross the same point in the diagram, we draw the string having lowest y-coordinate going over the other string. For simplicity, only two strings are allowed to cross at the same point.

Theorem 3.1. $\mathcal{B}_n \cong B_n$.

There is an intuitive mapping between braids on n strings and the Artin braid group. The twist between the ith and the $(i+1)$th string is mapped to either σ_i or σ_i^{-1}. It is however not immediately clear that this defines an isomorphism of groups. It was one of the accomplishments of Artin in [2] to establish the correspondence in Theorem 3.1.

3.3 Configuration spaces

In physics, the configuration space is the space of particle positions. Together with the space of momenta, it constitutes the phase space. Mathematically, it may be defined for any topological space. Particle exchange corresponds to continuous paths in configuration space, and so the structure of such paths, or rather loops, is of interest.

Definition 3.5. *Given a topological space M, the configuration space $F_n(M)$ of n ordered points is*

$$F_n(M) = \{(p_1, p_2, ..., p_n) \in M^n \mid p_i \neq p_j \text{ if } i \neq j\}$$

The configuration space $C_n(M)$ of n unordered points is

$$C_n(M) = F_n(M)/S_n = \{\{p_1, p_2, ..., p_n\} \in M^n \mid p_i \neq p_j \text{ if } i \neq j\}$$

where S_n is the symmetric group and $C_n(M)$ is given the quotient topology.

The configuration spaces may be interpreted as sets of distinct particles, and the difference between $F_n(M)$ and $C_n(M)$ is that of distinguishable and indistinguishable particles. As we are interested in applications to anyons, which are indistinguishable quasi-particles, it is ultimately $C_n(M)$ that is of interest in this thesis.

The structure of loops in a topological space is captured by the fundamental group. Before defining this, and examining the connection between configuration spaces, fundamental groups and the braid group, some preliminary definitions are needed. See [19], especially Chapter 4, for a more extensive review of these topics in connection to physics. For a standard reference in algebraic topology, see [8].

Definition 3.6. *A path f in a topological space M from $x_0 \in M$ to $x_1 \in M$ is a continuous function $I \to M$, with $I = [0,1]$, such that $f(0) = x_0$, $f(1) = x_1$.*

Definition 3.7. *A loop γ at x_0 in M is a path with $\gamma(0) = \gamma(1) = x_0$.*

Definition 3.8. *Two loops f and g at $x_0 \in M$ are said to be homotopic, written $f \simeq g$, if there exists a continuous map $F\colon I \times I \to M$ such that*

$$\begin{aligned} &F(t,0) = f(t) \\ &F(t,1) = g(t) && \forall t \in I \\ &F(0,s) = F(1,s) = x_0 && \forall s \in I \end{aligned}$$

The function F is called a homotopy from f to g.

The definition of homotopy defines an equivalence relation between loops, identifying loops that may be continuously deformed into one another as essentially the same. The set of all loops homotopic to a given loop f is called the homotopy class of f, denoted $[f]$. The loop f is called a representative of the homotopy class $[f]$.

Definition 3.9. *The fundamental group $\pi_1(M, x_0)$ of a topological space M at the base point $x_0 \in M$ is the set of homotopy classes of loops at x_0, with multiplication of classes $[f]$ and $[g]$ given by $[f] * [g] = [fg]$, where the multiplication of f and g is defined as*

$$(fg)(t) = \begin{cases} f(2t) & \text{if } 0 \le t \le 1/2 \\ g(2t-1) & \text{if } 1/2 \le t \le 1. \end{cases}$$

If the topological space M is path connected, i.e. any two points in M may be joined by a path, then up to isomorphism the fundamental group is independent of the base point x_0, which may then be omitted from the notation, writing only $\pi_1(M)$.

The next theorem draws the important connection between braids and anyons. Although the proof is rather technical, the correspondence between braids and loops in the configuration space of particles in the plane is rather intuitive; imagining a braid as the evolution in time as particles in the plane exchange places, as in Figure 3.5, a loop in $C_n(\mathbb{R}^2, p)$ is simply the evolution of the particles traversing the braid projected onto the plane.

Theorem 3.2. *$B_n \cong \pi_1(C_n(\mathbb{R}^2, p))$ where $p = \{(1,0), (2,0), ..., (n,0)\}$.*

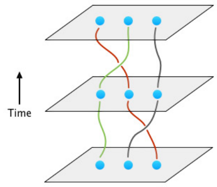


Figure 3.5. Particles in the plane form a braid in space time when changing place

Proof. An isomorphism between $\mathcal{B}_n$ and $\pi_1(C_n(\mathbb{R}^2))$ is obtained by first defining the map φ sending a geometric braid b to the loop $\varphi(b)\colon I \to C_n(\mathbb{R}^2)$ where $t \mapsto \{r_1(t), r_2(t), ..., r_n(t)\}$ and $r_i(t)$ is the intersection of the ith string b_i of b with $\mathbb{R}^2 \times \{t\}$ projected onto $\mathbb{R}^2$. By the properties of geometric braids, it is clear that φ is indeed a loop. Specifically, φ is continuous. We will see that φ induces a bijection $\tilde{\varphi}$ from isotopy classes of geometric braids to $\pi_1(C_n(\mathbb{R}^2))$.

We first check that $\tilde{\varphi}$ is well-defined, i.e. that isotopic braids map to the same homotopy class. Let b, b' be isotopic geometric braids, and let $F\colon b \times I \to \mathbb{R}^2 \times I$ be an isotopy from b to b'. For $\tilde{\varphi}$ to be well-defined, we need to show that $\varphi(b) \simeq \varphi(b')$. A homotopy is constructed from $\varphi(b)$ to $\varphi(b')$ as

$$\begin{array}{rccc} H\colon & I \times I & \to & C_n(\mathbb{R}^2) \\ & (t,s) & \mapsto & \varphi\left(F(b,s)\right)(t) \end{array}$$

Continuity of H follow from the continuity of F and φ. Furthermore, since F is an isotopy of braids and hence $F(b,s)$ is a geometric braid $\forall s \in I$, we have that, $\forall t \in I$

$$\begin{array}{ll} H(t,0) = \varphi\left(F(b,0)\right)(t) = \varphi\left(b\right)(t) & H(0,s) = \varphi\left(F(b,s)\right)(0) = p \\ H(t,1) = \varphi\left(F(b,1)\right)(t) = \varphi\left(b'\right)(t) & H(1,s) = \varphi\left(F(b,s)\right)(1) = p \end{array}$$

Hence, H is a homotopy, and $\varphi(b) \simeq \varphi(b')$.

$\tilde{\varphi}$ is injective: Assume that $\varphi(b) \simeq \varphi(b')$ for geometric braids b, b', and let H be a homotopy from $\varphi(b)$ to $\varphi(b')$. Every loop in $C_n(\mathbb{R}^2, p)$ induces a set of paths in $\mathbb{R}^2$, starting and ending at distinct pair of points in p. Thus every homotopy between loops in $C_n(\mathbb{R}^2, p)$ induces a set of homotopies between paths in $\mathbb{R}^2$. Label the homotopies of paths corresponding to H as H_i for $i = 1, ..., n$, where H_i is the unique homotopy from the path starting at $(i, 0) \in p$. Now, define an isotopy F from b to b' by

$$\begin{array}{rccc} F\colon & b \times I & \to & \mathbb{R}^2 \times I \\ & (x,s) & \mapsto & (H_i(t,s), t) \end{array}$$

where i is the index of the string $b_i \subset b$ containing x, and x is uniquely written $x = (r_i(t), t)$ for some $t \in I$.

Checking that F is an isotopy, first note that continuity follows from continuity of each H_i and the continuity of t when varying x. As F is continuous, b is compact and $\mathbb{R}^2 \times I$ is Hausdorff, F_s is an embedding if F_s is injective. To see this, take $x, x' \in b$, $x \neq x'$. We have $x = (r_i(t), t)$, $x' = (r_j(t'), t')$ for $i \neq j$ or $t \neq t'$, or both. As $H_i(t, s) \neq H_j(t', s)$ $\forall t \in I$ when $i \neq j$ and $F_s(x) = (H_i(t, s), t)$, $F_s(x') = (H_j(t', s), t')$ it is clear that $F_s(x) \neq F_s(x')$. Hence F_s is an embedding. From the continuity of the H_i:s and the definition of F, it is clear that $F_s(b)$ is a geometric braid. Hence F is an isotopy, and $\tilde{\varphi}$ is injective.

$\tilde{\varphi}$ is surjective: Surjectivity of φ follows by assigning to an arbitrary loop γ in $C_n(\mathbb{R}^2, p)$ the geometric braid $b = \bigcup_{t \in I} \gamma(t) \times \{t\}$. It is clear that $\varphi(b) = \gamma \implies \tilde{\varphi}$ is surjective, showing $\mathcal{B}_n \cong \pi_1 \left(C_n(\mathbb{R}^2, p)\right)$.

According to Theorem 3.1, $B_n \cong \mathcal{B}_n$ and thus $B_n \cong \pi_1 \left(C_n(\mathbb{R}^2, p)\right)$. □

Considering instead $F_n(\mathbb{R}^2)$ would yield the pure braid group P_n.

The next lemma may be used to find the fundamental group for dimensions greater than two. It is a standard result in algebraic topology, and it appears for example in [1] as Theorem 5.13.

Lemma 3.1. *If X is a simply connected topological space, G is a discrete group acting continuously on X, and $\forall x \in X$ $\exists U \subset X$ open neighbourhood of x such that $U \cap g(U) = \varnothing$ $\forall g \in G$, then $\pi_1 (X/G) \cong G$.*

Theorem 3.3. *$\pi_1 (C_n(\mathbb{R}^m)) \cong S_n$ for $m \geq 3$.*

Proof. Let $m \geq 3$. $F_n(\mathbb{R}^m)$ is simply connected, and the action of S_n by permuting indices is continuous. For a point $p \in F_n(\mathbb{R}^m)$ and an open neighbourhood U of $p = (p_1, p_2, ..., p_n)$, we can choose disjoint open neighbourhoods $U_1, U_2, ..., U_n$ in $\mathbb{R}^m$ of $p_1, p_2, ..., p_n$ respectively with $\text{diam}(U_k) < \frac{1}{2} \min(d(p_i, p_j))$ for $i, j, k \in \{1, 2, ..., n\}$. It is clear that with $U = U_1 \times U_2 \times ... \times U_n$ as the open neighbourhood of p, the requirement of Lemma 3.1 is satisfied, and thus $\pi_1 (C_n(\mathbb{R}^m)) \cong S_n$. □

3.4 The center of the braid group

Characterising those elements of the braid group that commute with all other elements will be useful when investigating representations. In group theory, this is known as the center of a group, denoted with Z.

Theorem 3.4. *For $n \geq 3$, $Z(B_n) = \langle \theta_n \rangle$ where $\theta_n = \Delta_n^2$ and*

$$\Delta_n = (\sigma_1 \sigma_2 \cdots \sigma_{n-1})(\sigma_1 \sigma_2 \cdots \sigma_{n-2}) \cdots (\sigma_1 \sigma_2)\sigma_1.$$

This is proven in [9], a part of which is presented here. To see that θ_n is in $Z(B_n)$, we use the visual aid of geometric braids. Firstly, Δ_n is obtained from the trivial

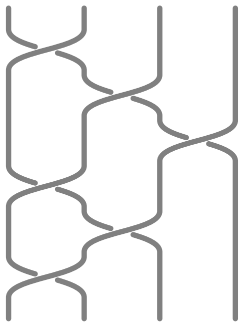


Figure 3.6. Braid diagram of Δ_4.

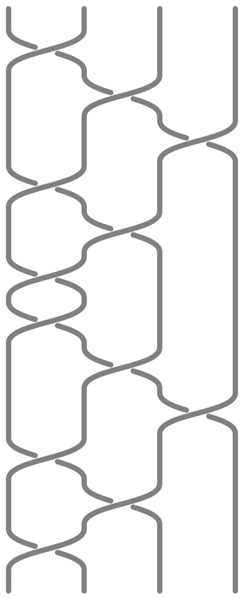


Figure 3.7. Braid diagram of θ_4.

braid by twisting the bottom set of points by π while keeping the top fixed, see Figure 3.6. It is therefore sometimes called the half twist. θ_n is given similarly by a full twist of 2π, and is called the full twist, shown in Figure 3.7. From this geometric description, it is clear that $\theta_n \in P_n$. For every generator σ_i we have $\sigma_i\Delta_n = \Delta_n\sigma_{n-i}$, cf. Figure 3.8, from which it follows that

$$\sigma_i\theta_n = \sigma_i\Delta_n\Delta_n = \Delta_n\sigma_{n-i}\Delta_n = \Delta_n\Delta_n\sigma_i = \theta_n\sigma_i.$$

Hence the full twist commutes with every generator, and so $\theta_n \in Z(B_n)$. It follows that if $Z(P_n) = \langle\theta_n\rangle$ then $Z(P_n) \subset Z(B_n)$. Conversely,

$$\pi(Z(B_n)) \subset Z(S_n) = \{1\} \implies Z(B_n) \subset Z(P_n)$$

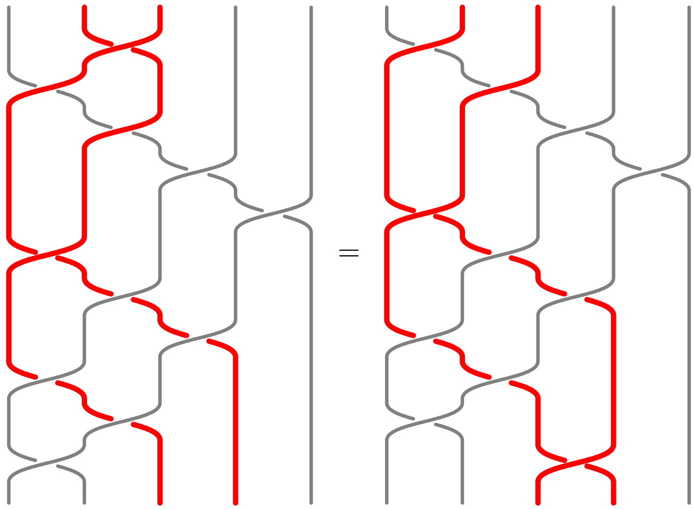


Figure 3.8. $\sigma_2\Delta_5 = \Delta_5\sigma_3$

and hence $Z(B_n) = Z(P_n)$, where we have used the known fact that $Z(S_n)$ is trivial for $n \geq 3$. To prove the theorem it suffices then to prove that $Z(P_n) = \langle\theta_n\rangle$. This is done by induction in [9], Theorem 1.24. Note that θ_n may be expressed inductively as $\iota(\theta_{n-1})\gamma_n$ where ι is the inclusion map $P_{n-1} \to P_n$ and

$$\gamma_n = \sigma_{n-1}\sigma_{n-2}\cdots\sigma_1\sigma_1\cdots\sigma_{n-2}\sigma_{n-1}.$$

This is seen readily by comparing Figure 3.7 and Figure 3.9.

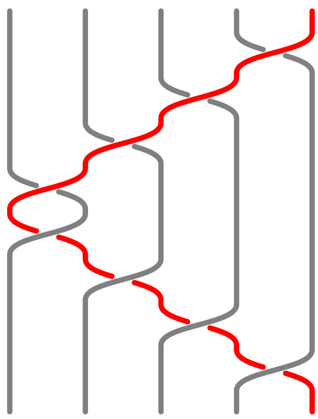


Figure 3.9. Braid diagram showing γ_5. Note that the restriction to the first four strings is the identity.

3.5 Alternative presentations

As previously mentioned, there are other presentations of the braid group than the Artin presentation. This section presents one such presentation, with particularly few generators.

Theorem 3.5. *For any n, B_n may be presented with at most two generators. More specifically, for $n \geq 3$, B_n is generated by σ_1 and $\alpha = \sigma_1\sigma_2 \cdots \sigma_{n-1}$.*

Proof. For $n = 1$ and $n = 2$, there is nothing to prove, so let $n \geq 3$. The proposition follows from the relation

$$\sigma_i = \alpha^{i-1}\sigma_1\alpha^{1-i} \quad \forall\ i = 1, ..., n-1$$

which we prove by induction. Let $k = 1, ..., i$. For $k = 1$ it is clear that $\alpha^{i-1}\sigma_1\alpha^{1-i} = \alpha^{i-k}\sigma_k\alpha^{k-i}$. Assume this relation is valid for $k-1$. Then

$$\begin{aligned} \alpha^{i-1}\sigma_1\alpha^{1-i} = \alpha^{i-(k-1)}\sigma_{k-1}\alpha^{k-1-i} &= \alpha^{i-k}\alpha\sigma_{k-1}\alpha^{-1}\alpha^{k-i} \\ &= \alpha^{i-k}\sigma_k\alpha^{k-i} \end{aligned}$$

since

$$\begin{aligned} \alpha\sigma_{k-1}\alpha^{-1} &= \sigma_1 \cdots \sigma_{k-2}\sigma_{k-1}\sigma_k\sigma_{k-1}\sigma_{k+1} \cdots \sigma_{n-1}\alpha^{-1} = \\ &= \sigma_1 \cdots \sigma_{k-2}\sigma_k\sigma_{k-1}\sigma_k\sigma_{k+1} \cdots \sigma_{n-1}\alpha^{-1} = \\ &= \sigma_k\sigma_1 \cdots \sigma_{k-2}\sigma_{k-1}\sigma_k\sigma_{k+1} \cdots \sigma_{n-1}\alpha^{-1} = \\ &= \sigma_k\alpha\alpha^{-1} = \sigma_k \end{aligned}$$

by the braid relations. Hence

$$\alpha^{i-1}\sigma_1\alpha^{1-i} = \alpha^{i-k}\sigma_k\alpha^{k-i} \quad \forall\ k = 1, ..., i$$

from which the proposition follows. ☐

This means that no matter the size of the braid group, all information is in a sense contained in only two elements. Although the above presentation is simpler than the standard presentation in the sense that the generators are fewer, the relations are more complicated, and for $n > 3$ they are rather unwieldy to work with. In general they are given by

$$\begin{aligned} \sigma_1\alpha\sigma_1\alpha^{-1}\sigma_1 &= \alpha\sigma_1\alpha^{-1}\sigma_1\alpha\sigma_1\alpha^{-1} \\ \alpha^{i-j}\sigma_1\alpha^{j-i}\sigma_1 &= \sigma_1\alpha^{i-j}\sigma_1\alpha^{j-i} \qquad \text{for } |i-j| \geq 2. \end{aligned}$$

Furthermore, the formula $\sigma_i = \alpha^{i-1}\sigma_1\alpha^{1-i}$ immediately imply that the standard generators of the braid group are conjugate.

Corollary 3.1. *For any $n \in \mathbb{N}$, the standard generators of B_n are conjugate.*

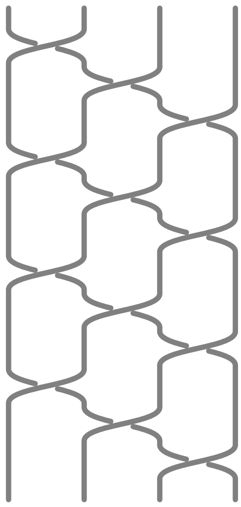


Figure 3.10. Braid diagram of α_4^4.

There is a connection between the generator of the center of B_n, i.e. the full twist θ, and the element α, both corresponding to n strings.

Proposition 3.2. *$\theta = \alpha^n$.*

Proof. Note that the permutation induced by α_n is $p(\alpha) = (n, 1, ..., n-1) \in S_n$, and hence $p(\alpha^n) = (1, ..., n)$ and $\alpha^n \in P_n$. The first non-trivial case is $n = 3$, for which the proposition is just a direct application of the second braid relation: $\theta_3 = \sigma_1\sigma_2\sigma_1\sigma_1\sigma_2\sigma_1 = \sigma_1\sigma_2\sigma_1\sigma_2\sigma_1\sigma_2 = \alpha_3^3$. Assume now that $\alpha_{n-1}^{n-1} = \theta_{n-1}$. From Figure 3.9 and Figure 3.10, it is clear that $\alpha_n^n = \iota(\alpha_{n-1}^{n-1})\gamma_n = \iota(\theta_{n-1})\gamma_n = \theta_n$ by the induction hypothesis. □

This especially means that α^n generate the center of B_n.

Setting $\Omega = \sigma_1\alpha$, we obtain a variant of the presentation above, with Ω and α as generators. The relations from before now become

$$\begin{aligned} \Omega^2\alpha^{-2}\Omega &= \alpha\Omega\alpha^{-2}\Omega^2\alpha^{-2} \\ \alpha^{i-j}\Omega\alpha^{j-i-1}\Omega &= \Omega\alpha^{i-j-1}\Omega\alpha^{j-i} \quad \text{for } |i-j| \geq 2. \end{aligned}$$

However, Ω and α also satisfy the simpler relation

$$\Omega^{n-1} = \alpha^n \ .$$

This may be seen by noting that $\sigma_i = \alpha^{i-1}\sigma_1\alpha^{1-i}$ implies $\sigma_{i+1} = \alpha\sigma_i\alpha^{-1}$, giving in turn $\alpha\sigma_k = \sigma_{k+1}\alpha$. By repeated use of this relation and explicitly writing out Ω^{n-1}, the relation between Ω and α then follows.

For $n = 3$, the relation $\Omega^{n-1} = \alpha^n$ is in fact a sufficient relation for generating the braid group as in this case Ω^2 lies in the center of B_3 and the other relations become trivial.

Proposition 3.3. $B_3 = \langle \alpha, \Omega \mid \Omega^2 = \alpha^3 \rangle$.

Chapter 4

Representations of the braid group

4.1 Definitions and properties

For infinite discrete groups such as the braid group, the techniques and results from the theory of representations of finite groups, such as character theory, and the extensions to compact topological groups and Lie groups does not apply. In this section, apart from the basic definitions, some properties of representations applying to braid groups are examined. For further details on the basic definitions for representations, see Chapter 11 in [6].

Definition 4.1. *A (linear) representation ρ of a group G is a group homomorphism $\rho\colon G \to GL(V)$ where V is a vector space. The representation ρ is then said to be a representation on V.*

In this thesis, the interest is towards finite dimensional complex vector spaces, and we henceforth denote the general linear group of $\mathbb{C}^n$ by $GL_n(\mathbb{C})$. If not otherwise stated, every vector space is implicitly understood to be complex.

Every representation induces a group action of G on V, and conversely, every group action defines a corresponding representation. Thus representations of a group G on a vector space V and group actions of G on V are one and the same. In order to define a $\mathbb{C}G$-module V from a vector space V, it is sufficient to define the action of the group elements on V. Furthermore, since every $\mathbb{C}G$-module defines a group action, we see that representations of G on V and modules V over $\mathbb{C}G$ are in one-to-one correspondence.

Definition 4.2. *The degree of a representation on a vector space V is the dimension of V.*

Because of this definition, the degree of a representation ρ is often called instead the dimension of ρ, and the two terms will be used interchangeably when there exists no risk of confusion. For example, Section 4.5 refers to low-dimensional representations.

Definition 4.3. *The matrix algebra of a complex representation ρ of a group G is the $\mathbb{C}$-algebra $\rho(\mathbb{C}G)$ of matrices where $\mathbb{C}G$ is the group algebra and ρ is extended linearly.*

The generic matrix algebra will be denoted $\mathcal{A}$. It is the algebra generated via matrix addition and multiplication and scalar multiplication from the image of the group generators. As any $\mathbb{C}$-algebra, $\mathcal{A}$ is also a vector space over $\mathbb{C}$, given as $\text{span}\{\rho(g)\}$, $g \in G$.

Definition 4.4. *Two representations ρ and τ of a group G are similar, written $\rho \cong \tau$, if there exists an invertible matrix P such that*

$$\rho(g)P = P\tau(g)$$

$\forall g \in G$.

Often the group elements are omitted from the notation, writing $\rho P = P\tau$. The representation ρ is similar to τ if $\rho = P\tau P^{-1}$, and so similarity between representations expresses simultaneous conjugation between elements in the image of the representations.

Definition 4.5. *A representation ρ on a vector space V is said to be reducible if there is a proper subspace $M \subset V$ that is invariant under the group action induced by ρ, i.e. $\rho(g)m \in M \ \forall g \in G,\ m \in M$. Equivalently, ρ is reducible if there is a proper submodule $M \subset V$ of the $\mathbb{C}G$-module V induced by ρ.*

Definition 4.6. *If a representation is not reducible, it is called irreducible, or alternatively, simple, the latter terminology being more common when referring to modules.*

Definition 4.7. *Let ρ be a reducible representation and U an invariant subspace. The subrepresentation $\hat{\rho}$ of ρ on U is the representation on U obtained by restricting ρ to U.*

By this definition, it is clear that subrepresentations are exactly the proper submodules of the associated group algebra module of the representation.

Definition 4.8. *A representation on a vector space V is called completely reducible, or semisimple, if V can be decomposed into the direct sum of finitely many invariant proper subspaces.*

Equivalently, the representation is semisimple if the $\mathbb{C}G$-module V can be decomposed as the direct sum of proper submodules. Sometimes however, the terms are taken to a priori refer to different, but related, properties. Semisimpleness is then defined as above, while complete reducibility is the property that given any proper submodule, there exists some other proper submodule such that their direct sum is the entire module. In the finite-dimensional setting however, these properties coincide and we need not make the distinction. The following result may be found with some additions in [9], and only the proof of one direction is presented here.

Proposition 4.1. *For a finite-dimensional $\mathbb{C}G$-module M, the following are equivalent*

(i) *There exists a finite number of proper simple submodules $\{M_i\}_{i\in I}$ of M such that $M = \bigoplus_{i\in I} M_i$*

(ii) *For any proper submodule $M' \subset M$, there exists a proper submodule $M'' \subset M$ such that $M = M' \oplus M''$.*

Proof. $(i) \implies (ii)$: Let $M = \bigoplus_{i\in I} M_i$ for a finite indexing set I, and assume that $M' \subset M$ is a proper submodule. There is a (possibly not unique) finite maximal subset $I' \subset I$ such that the sum $M' + \sum_{i\in I'} M_i$ is direct. Take M_k for $k \in I \setminus I'$. If $M_k \subset M' + \sum_{i\in I'} M_i$, then $M = M' + \sum_{i\in I'} M_i$ and we are done. Since $M_k \cap (M' + \sum_{i\in I'} M_i)$ is a submodule of the simple module M_k, either $M_k \cap (M' + \sum_{i\in I'} M_i) = M_k$, i.e. $M_k \subset M' + \sum_{i\in I'} M_i$, or $M_k \cap (M' + \sum_{i\in I'} M_i) = 0$. Assuming the latter, the sum $M' + \sum_{i\in I'} M_i + M_k$ is direct, contradicting the maximality of I'. Hence $M = M' \oplus M''$ with $M'' = \sum_{i\in I'} M_i$. □

The next theorem is known as Burnside's irreducibility criterion, stated here only in the complex case. A proof by elementary linear algebra due to [14] is given here. First, two related definitions and a lemma is given.

Definition 4.9. *The dual representation (also called contragredient representation) ρ' of a representation ρ on the vector space V is a representation on the dual vector space V' defined as*

$$\rho'(g) = \left(\rho(g^{-1})\right)^T \quad \forall g \in G$$

where G is the group and T denotes transpose.

The use of both inverse and transpose in the above definition ensures that ρ' is indeed a homomorphism from G and hence a representation.

Definition 4.10. *A representation ρ on a vector space V is called transitive if $\forall x \in V \setminus \{0\}$, $\{Ax \mid A \in \mathcal{A}\} = V$ where $\mathcal{A}$ is the associated matrix algebra.*

Lemma 4.1. *A representation of finite dimension is irreducible if and only if it is transitive.*

Proof. Let $\mathcal{A}$ be the associated matrix algebra. Then $\forall x \in V, x \neq 0$, $\{Ax \mid A \in \mathcal{A}\}$ is an invariant subspace of V. Hence the representation is irreducible if and only if it is transitive. □

Theorem 4.1. *Burnside's irreducibility criterion.*
Denote with $M_n(\mathbb{C})$ the algebra of all $n \times n$ complex matrices. Let ρ be a representation on $\mathbb{C}^n$ of a group G, with associated matrix algebra $\mathcal{A}$. Then ρ is irreducible if and only if $\mathcal{A} = M_n(\mathbb{C})$.

Proof. We prove by induction on n that ρ transitive implies $\mathcal{A} = M_n(\mathbb{C})$. This will prove the theorem, since by Lemma 4.1, transitivity and irreducibility are equivalent. Furthermore, all matrices are linear combinations of rank 1 matrices, so it will suffice to show that the algebra contains all rank 1 transformations on $\mathbb{C}^n$.

For $n = 1$ the vector space is $\mathbb{C}$ and the representation acts by multiplication. It is clear that transitivity implies $\mathcal{A} = M_1(\mathbb{C}) = \mathbb{C}$.

Now, assume that the statement is valid for $n - 1$, and let ρ be transitive of dimension n. There exists a non zero transformation $T \in \mathcal{A}$ that is not a scalar multiple of the identity. Either T is singular, or we can form the transformation $\lambda T - T^2$, where λ is an eigenvalue of T. This transformation is then singular, since for an eigenvector v, $\left(\lambda T - T^2\right) v = 0$. This argument shows that there is a singular transformation $F \in \mathcal{A}$.

For such a transformation F, the restriction $F\mathcal{A}$ of $\mathcal{A}$ to the vector space $F\mathbb{C}^n$ is a transitive algebra of dimension less than n. By the induction hypothesis, this algebra is then the entire algebra of linear transformations on $F\mathbb{C}^n$. Especially, there exists an element $A_0 \in \mathcal{A}$ such that FA_0 is a transformation of rank 1 on $F\mathbb{C}^n$, and hence FA_0F is a transformation of rank 1 on $\mathbb{C}^n$.

Every rank 1 transformation in $\mathcal{A}$ is of the form $\mathbb{C}^n \ni x \mapsto vf^Tx$ where $v, f \in \mathbb{C}^n$. Let $T_0 \in \mathcal{A}$ be of rank 1, mapping $x \mapsto v_0f_0^Tx$. For $A \in \mathcal{A}$, we have $AT_0x = Av_0f_0^Tx$ and $T_0Ax = v_0f_0^TAx = (A^Tf_0)^Txv_0$ $\forall x \in \mathbb{C}^n$. Note that A^T is an element in the dual algebra $\mathcal{A}'$ of $\mathcal{A}$. It is clear that transitivity of $\mathcal{A}$ implies transitivity of $\mathcal{A}'$, and so every transformation of rank 1 may be generated from T_0. Hence $\mathcal{A}$ contains all rank 1 transformations in $M_n(\mathbb{C})$, and thus $\mathcal{A} = \mathbb{C}^n$. □

Calling a matrix scalar if it is a scalar times the identity, we have the following well-known result.

Corollary 4.1. *Schur´s first lemma.*
A matrix commuting with an irreducible representation is scalar.

Proof. Denote the matrix by M. By Theorem 4.1, M commutes with every matrix, and hence must be a scalar matrix. □

In the following, a representation mapping all group elements to the same scalar matrix will be called trivial.

Proposition 4.2. *If a representation ρ of B_n maps any of the generators σ_i to a scalar matrix, say dI for $d \in \mathbb{C}$, then ρ is trivial.*

Proof. Assume $\sigma_i \mapsto dI$. Then the braid relations imply

$$\rho(\sigma_i)\rho(\sigma_{i+1})\rho(\sigma_i) = \rho(\sigma_{i+1})\rho(\sigma_i)\rho(\sigma_{i+1}) \implies d^2\rho(\sigma_{i+1}) = d\rho(\sigma_{i+1})^2 \implies \rho(\sigma_{i+1}) = dI$$

and similarly for $\rho(\sigma_{i-1})$. □

Before giving the next theorem, we make the following definition.

Definition 4.11. *An extension $\tilde{\rho}$ of a representation ρ of B_n to B_{n+k} is a representation of B_{n+k} such that $\tilde{\rho}(\sigma_i) = \rho(\sigma_i)$ for $i = 1, ..., n-1$.*

By Theorem 4.1, it is clear that any extension of an irreducible representation is irreducible.

Theorem 4.2. *An irreducible representation of B_n can not be extended to a non-trivial representation of B_{n+2}.*

Proof. By the braid relations, the image of σ_{n+1} must commute with the matrix algebra of B_n. By Burnside's irreducibility criterion, this consists of all matrices, and thus the image of σ_{n+1} is scalar. By Proposition 4.2, all generators must then be mapped to this scalar matrix and hence the extension is trivial. □

4.2 Bosons, fermions and abelian anyons

It is the results of Theorems 3.2 and 3.3, with the occurrence of the braid group instead of the symmetric group for $\mathbb{R}^2$, that constitute the essential difference between particle statistics in two and three dimensions. In this section we investigate the possible unitary representations of degree one in the two cases. For two dimensions and higher, there are two such representations, corresponding to bosons and fermions, while for two dimensions we get abelian anyons.

Let $\rho\colon G \to \mathbb{U}(1)$ be a representation, where G is either S_n or B_n. Since ρ is complex and of degree one, all elements in the image of ρ will commute. In both groups, the relation $\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}$ for $0 \leq i \leq n-2$, with σ_i being the generators, implies

$$\begin{aligned} \rho(\sigma_i)\rho(\sigma_{i+1})\rho(\sigma_i) &= \rho(\sigma_{i+1})\rho(\sigma_i)\rho(\sigma_{i+1}) \\ \rho(\sigma_{i+1})\rho(\sigma_i)\rho(\sigma_i) &= \rho(\sigma_{i+1})\rho(\sigma_{i+1})\rho(\sigma_i) \\ \rho(\sigma_i) &= \rho(\sigma_{i+1}) \end{aligned}$$

Hence ρ maps every element in G to the same unitary complex number, say $z = \exp(i\theta)$.

For $G = S_n$, there is the additional relation $\sigma_i^2 = e$, implying $\exp(i2\theta) = 1$ and $\theta = 0$ or π, giving $z = 1$ or $z = -1$. The representation ρ gives the phase multiplying the wave function under particle exchange, and so this result means that in three dimensions and higher, there are exactly two types of wave functions - one symmetric and one antisymmetric under exchange of particles. This corresponds to two types of fundamental particles - bosons and fermions.

For $G = B_n$, the situation is different. There is no extra relation forcing θ to take on specific values, and we are left with a family of possible representations, with $\theta \in [0, 2\pi)$. As mentioned in the introduction, particles realising such representations are called abelian anyons.

4.3 Representations of B_3 in $SU(2)$

For wave functions in higher dimension, and thus representations of higher degree, the elements in the image do not necessarily commute, and the possibilities are richer. Particles realising such representations are called non-abelian anyons. In this section, following [10], the special case of representations of B_3 in $SU(2)$ using its connection to the quaternions is studied. This may then be extended to representations in $U(2)$ in a straightforward manner.

The elements of $SU(2)$ are complex 2×2 matrices $\begin{pmatrix} z & w \\ x & y \end{pmatrix}$, $x, y, z, w \in \mathbb{C}$, fulfilling the requirements $SS^* = S^*S = I$ and $\det(S) = 1$. This gives a set of equations for x, y, z, w which after some tedious manipulations show the following.

Proposition 4.3. *Every $S \in SU(2)$ is of the form*

$$S = \begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix} \quad z, w \in \mathbb{C}, \ |z|^2 + |w|^2 = 1$$

where $\bar{z}$ denotes the complex conjugate of z.

The key element in the treatment is now to note that $SU(2)$ may be identified with the real algebra generated by the quaternions. This can be seen by first writing $z = a + bi$, $w = c + di$ for $a, b, c, d \in \mathbb{R}$, in which case it follows from the above that an element S of $SU(2)$ takes the form

$$S = a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} + c \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

Writing

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \mathbf{i} = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} \quad \mathbf{j} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \mathbf{k} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

and noting that $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -\mathbf{1}$ and $\mathbf{ij} = \mathbf{k}$, $\mathbf{jk} = \mathbf{i}$, $\mathbf{ki} = \mathbf{j}$ with $\mathbf{i}, \mathbf{j}, \mathbf{k}$ anti-commuting with each other, we see that S may indeed be identified with a linear combination of quaternions.

A quaternion q is called pure if its real part is zero, and it is called a unit quaternion if its norm $\sqrt{qq^*}$ is 1, where here $*$ is quaternion conjugation, corresponding to changing sign of all non-real components. This means that $S = a + bu$, where

$a, b \in \mathbb{R}$ and u is a pure unit quaternion, and $a^2 + b^2 = 1$ since S has determinant 1. As a pure unit quaternion has three real, linearly independent components, we can identify them with elements in $\mathbb{R}^3$, and because of their unit norm, more specifically with S^2. With this remark in mind, and from the quaternion relations, one calculates that any pure unit quaternion squares to negative 1, and that for two pure unit quaternions u and v, $uv = -u \cdot v + u \times v$.

We now look at a representation ρ of B_3 in $SU(2)$. This means describing quaternions to the braid generators. Let $\rho(\sigma_1) = g$ and $\rho(\sigma_2) = h$ be quaternions. The braid relations give $ghg = hgh$, or equivalently $h^{-1}gh = ghg^{-1}$. Letting $g = a + bu$ and $h = c + dv$, with $a^2 + b^2 = c^2 + d^2 = 1$ and u, v pure unit quaternions, we have for a pure quaternion P

$$gPg^{-1} = (a^2 - b^2)P + 2ab(P \times u) + 2(P \cdot u)b^2u.$$

The expression gPg^{-1} has the geometric interpretation of a rotation of the vector P around the u direction. The braid relations imply that $c + dgvg^{-1} = a + bh^{-1}uh$. Using the above expression for a rotation

$$\begin{gathered} c + d\left[(a^2 - b^2)v + 2ab(v \times u) + 2(v \cdot u)b^2u\right] = \\ a + b\left[(c^2 - d^2)u - 2cd(u \times v) + 2(v \cdot u)d^2v\right] \end{gathered}$$

and thus the equations

$$\begin{gathered} c = a \\ d(a^2 - b^2) = 2(v \cdot u)d^2b \\ b(c^2 - d^2) = 2(v \cdot u)b^2d \\ 2abd - 2bcd = 0 \end{gathered}$$

where it has been assumed that $u \neq v$.

From these equations, there are three possibilities. b and d may both be zero, in which case $g = a$, $h = c$ with a and c both either 1 or -1. This corresponds to scalar matrices, with the generators in B_3 all mapping to plus or minus the identity. If only one of b and d are zero, we get either $g = 0$ or $h = 0$, and hence this case is excluded. If both b and d are non-zero, it follows that $d = \pm b$, in which case we also have $a^2 - \frac{1}{2} = \pm(v \cdot u)b^2$. Of course, $-v$ is a pure unit quaternion if v is, and we can work with $d = b$ and $-v$ instead. This discussion is summarized in a theorem, cf. [10].

Theorem 4.3. *All representations of B_3 in $SU(2)$, except the one mapping each generator to plus or minus the identity, is of the following form, which is given in quaternion form*

$$\rho(\sigma_1) = a + bu, \quad \rho(\sigma_2) = a + bv, \quad a^2 - \tfrac{1}{2} = (v \cdot u)b^2, \quad a^2 + b^2 = 1$$

for some pure unit quaternions u and v.

It is possible to extend any representation of B_3 in $SU(2)$ to a representation of B_4 simply by setting $\rho(\sigma_3) = \rho(\sigma_1)$. In fact, we shall see in Section 4.5 that this is true for any representation of B_3. However, using Theorem 4.2, it is not possible to extend further to B_5.

Although Theorem 4.3 is concerned with special unitary representations, it generalizes to unitary representations as every element $U \in U(2)$ may be written $U = zS$ for some $S \in SU(2)$ and $z \in \mathbb{C}$ with $|z| = 1$. This may be stated as $U(2) = \mathbb{T} \otimes SU(2)$ with $\mathbb{T}$ being the complex unit circle. Defining the representation $\chi(y)$ sending all generators in B_3 to $y \in \mathbb{C}^*$, any representation τ of B_3 in $U(2)$ may be written $\tau = \chi(y) \otimes \rho$ for some $y \in \mathbb{C}^*$ and representation ρ of B_3 in $SU(2)$, with the requirement $|y| = 1$.

4.4 The Burau representation

The Burau representation is probably the most commonly known non-abelian representation of the braid group, and was first presented in 1936 by W. Burau in [5]. It was the first interesting representation of the braid group to be studied. It is given here in three forms, of which the two reduced forms will be of greatest interest to us in the rest of the thesis.

In the following, let $\mathbb{C}^*$ denote the set of invertible elements in $\mathbb{C}$, i.e. $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$.

Definition 4.12. *The full Burau representation* $\hat{\beta}\colon B_n \to GL_n(\mathbb{C})$ *for* $n \geq 2$ *is given in matrix form by*

$$\hat{\beta}\left(\sigma_i\right) = I_{i-1} \oplus \begin{pmatrix} 1-z & z \\ 1 & 0 \end{pmatrix} \oplus I_{n-i-1} \quad 1 \leq i \leq n-1, \; z \in \mathbb{C}^*$$

Note that the full Burau representation is reducible, since for reducibility it suffices to check that the image of the group generators fixes a proper subspace of $\mathbb{C}^n$, and the subspace $\operatorname{span}\{e_1 + ... + e_n\}$ is invariant under $\hat{\beta}\left(\sigma_i\right), \; 1 \leq i \leq n-1$. This is stated more explicit in the next proposition.

Proposition 4.4. *The full Burau representation is reducible, since for*

$$C = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \quad \textit{we have} \quad C^{-1}\hat{\beta}(\sigma_i)C = \begin{pmatrix} \beta(\sigma_i) & 0 \\ \star_i & 1 \end{pmatrix}$$

where

$$\beta(\sigma_1) = \begin{pmatrix} -z & 0 \\ 1 & 1 \end{pmatrix} \oplus I_{n-3}, \qquad \beta(\sigma_{n-1}) = I_{n-3} \oplus \begin{pmatrix} 1 & z \\ 0 & -z \end{pmatrix}$$

$$\beta(\sigma_i) = I_{i-2} \oplus \begin{pmatrix} 1 & z & 0 \\ 0 & -z & 0 \\ 0 & 1 & 1 \end{pmatrix} \oplus I_{n-i-2} \qquad 1 < i < n-1$$

and $\star_i$ *is the vector of length* $n-1$ *given by*

$$\begin{aligned} \star_i &= (0, ..., 0, 0) \quad \textit{if} \quad i = 1, ..., n-2 \\ \star_i &= (0, ..., 0, 1) \quad \textit{if} \quad i = n-1 \end{aligned}$$

The invariance of $\operatorname{span}\{e_1 + ... + e_n\}$ is clearly exhibited in the form of the change-of-basis matrix C. Furthermore, it follows from the braid relations satisfied by $\hat{\beta}$ that β is also a representation of the braid group. Note that β is a representation on the quotient space $\mathbb{C}^n / \operatorname{span}\{e_1 + ... + e_n\}$.

Definition 4.13. *The reduced Burau representation* $\beta\colon B_n \to GL_{n-1}(\mathbb{C})$ *is the representation given by* $\beta(\sigma_i)$ *in Proposition 4.13.*

In the characterisation of low dimensional representations presented in Section 4.5, the following matrices play a central role due to their connection with the reduced Burau representation.

Definition 4.14. *A matrix* X *is called a pseudoreflection if* $X - I$ *has rank 1.*

Note that every pseudoreflection $X \in M_n(\mathbb{C})$ can be written in the form $X = I - ab^T$ where $a, b \in \mathbb{C}^n$.

Proposition 4.5. *A pseudoreflection* $X = I - ab^T$ *is invertible if and only if* $b^T a \neq 1$.

Proof. We show that the eigenvalues of X are 1 with multiplicity $n-1$ and $1 - b^T a$ with multiplicity 1, which shows the proposition. Furthermore, to find the eigenvalues is to calculate $\det(\lambda I - X) = \det(\lambda I - I + ab^T) = \det((\lambda - 1)I + ab^T) = (-1)^n \det((1-\lambda)I - ab^T)$, and hence we obtain the eigenvalues of X by calculating the eigenvalues of ab^T and changing $\lambda \mapsto 1 - \lambda$.

As ab^T is singular, it has the eigenvalue 0. Taking any vector w such that $\bar{w}$ is in the orthogonal complement of b, we have $ab^T w = a\langle b, \overline{w}\rangle = 0$. The subspace of such w has dimension $n-1$, and because the algebraic multiplicity is greater than or equal to the geometric multiplicity, we conclude that ab^T has eigenvalue 0 with multiplicity at least $n-1$. Multiplying the vector a with ab^T gives $(ab^T)a = a(b^T a) = (b^T a)a$ which shows that ab^T has eigenvalue $b^T a$. Hence ab^T has eigenvalue 0 with multiplicity $n-1$ and $b^T a$ with multiplicity 1. For X this corresponds to the eigenvalues 1 and $1 - b^T a$ respectively. ∎

Proposition 4.6. *The matrices $\beta(\sigma_i)$ are pseudoreflections $\forall i = 1, ..., n-1$.*

Proof. Set $a_1 = (z+1, -1, 0, ..., 0)^T$, $a_{n-1} = (0, ..., 0, -z, z+1)^T$ and $a_i = (0, ..., 0, -z, z+1, -1, 0, ..., 0)^T$ centred at i for $i = 2, ..., n-2$. Let $b_1, ..., b_{n-1}$ be the standard basis for $\mathbb{C}^{n-1}$. Then $\beta_n(\sigma_i) = I - a_i b_i^T$, i.e. a pseudoreflection. □

To prove the next theorem, we take a result from [7], Theorem 5.

Lemma 4.2. *Let Γ be the directed graph with vertices $1, ..., n$ with an edge from i to j, $i \neq j$, if $b_i^T a_j \neq 0$. A subgroup of $GL_n(\mathbb{C})$ generated by n pseudoreflections $I - a_1 b_1^T, ..., I - a_n b_n^T$ is irreducible if and only if Γ contains a directed path from i to j for $i \neq j$ and the matrix $(b_i^T a_j)$ is invertible.*

Theorem 4.4. *The reduced Burau representation β_n with $z \in \mathbb{C}$ as parameter is irreducible if and only if $z^{n-1} + z^{n-2} + ... + z + 1 \neq 0$.*

Proof. The relevant matrix from Lemma 4.2 is

$$(b_i^T a_j) = \begin{pmatrix} z+1 & -z & 0 & \cdots & 0 & 0 \\ -1 & z+1 & -z & 0 & \cdots & 0 \\ 0 & -1 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & z & 0 \\ 0 & \cdots & 0 & -1 & z+1 & -z \\ 0 & 0 & \cdots & 0 & -1 & z+1 \end{pmatrix}$$

with a_i, b_i from Proposition 4.6. This shows that there is indeed a directed graph as required, and hence according to Lemma 4.2 β is irreducible if and only $(b_i^T a_j)$ is invertible. Denote by M_k the matrix $(b_i^T a_j)$ for $i = 1, ..., k$. Then by expansion along the first row of M_k it is seen that $\det(M_k) = (z+1)\det(M_{k-1}) - z\det(M_{k-2})$. As $\det(M_1) = z+1$ and $\det(M_2) = z^2 + z + 1$ this shows via induction that $\det(M_{n-1}) = \det((b_i^T a_j)) = z^{n-1} + z^{n-2} + ... + z + 1$, and the theorem follows. □

If however z is a root to $z^{n-1} + ... + z + 1$, we have the following irreducible component of β.

Definition 4.15. *For $n \geq 4$ and $z^{n-1} + ... + z + 1 = 0$, the <u>fully reduced</u> Burau representation $\beta_n^r(z)\colon B_n \to GL_{n-2}(\mathbb{C})$ is the representation given by*

$$\beta_n^r(z)(\sigma_i) = \beta_{n-1}(z)(\sigma_i) \quad \text{for } i = 1, ..., n-2$$
$$\beta_n^r(z)(\sigma_{n-1}) = I - ab^T$$

where $a = (0, ..., 0, z)^T \in \mathbb{C}^{n-2}$ and
$b = -z(1, z+1, z^2 + z + 1, ..., z^{n-3} + ... + z + 1)^T \in \mathbb{C}^{n-2}$.

To see that β^r is indeed a representation, recall that $\beta(\sigma_i)$ is a pseudoreflection and may be written $I - a_i b_i^T$ with a_i, b_i as in Proposition 4.6, and hence $\beta^r(\sigma_i)$ is a pseudoreflection for every i. It is then easily verified that $\beta_n^r(\sigma_{n-1})\beta_n^r(\sigma_i) =$

$\beta_n^r(\sigma_i)\beta_n^r(\sigma_{n-1})$ for $i = 1, ..., n-3$ is due to the fact that $b_i^T a = b^T a_i = 0$ for $i = 1, ..., n-3$. Hence the first relation is satisfied. For the second relation, it may be calculated that $\beta_n^r(\sigma_{n-2})\beta_n^r(\sigma_{n-1})\beta_n^r(\sigma_{n-2}) = \beta_n^r(\sigma_{n-1})\beta_n^r(\sigma_{n-2})\beta_n^r(\sigma_{n-1})$ i.e. $(I - a_{n-2}b_{n-2}^T)(I - ab^T)(I - a_{n-2}b_{n-2}^T) = (I - ab^T)(I - a_{n-2}b_{n-2}^T)(I - ab^T)$ if and only if

$$(1 - b^T a_{n-2})a_{n-2}b_{n-2}^T = (1 - b^T a_{n-2})ab^T$$

where use has been made of $b_{n-2}^T a_{n-2} = z + 1$ and $b_{n-2}^T a = z$. Hence the second relation is satisfied if $b^T a_{n-2} = 1$, which is indeed the case, making use of the fact that $z^{n-1} + ... + z + 1 = 0$.

Note that any root of $z^{n-1} + ... + z + 1 = 0$ is not a root of $z^{n-2} + ... + z + 1 = 0$. It follows that the fully reduced Burau representation is irreducible, since it is then an extension of an irreducible representation.

4.5 Low-dimensional representations

In this section, we present a characterisation of irreducible non-abelian representations of B_n of degree less than or equal to $n - 1$, due to Formanek in [7].

Recall the representation χ defined in the end of Section 4.3. In the following, χ is defined in the same way but regarded as a function $B_n \to \mathbb{C}^*$ for arbitrary n.

Theorem 4.5. *If ρ is an irreducible representation of B_n of degree greater than or equal to 2 and less than or equal to $n-1$ with $n \neq 4, 5, 6$, then it is of Burau type, i.e.*

(i) $\rho \cong \chi(y) \otimes \beta_n(z)\colon B_n \to GL_{n-1}(\mathbb{C})$ where $z^{n-1} + ... + z + 1 \neq 0$
For $n = 3$, $\chi(y) \otimes \beta_3^r(z) \cong \chi(-yz) \otimes \beta_3^r(z^{-1})$.

(ii) $\rho \cong \chi(y) \otimes \beta_n^r(z)\colon B_n \to GL_{n-2}(\mathbb{C})$ where $z^{n-1} + ... + z + 1 = 0$
For $n = 4$, $\chi(y) \otimes \beta_4^r(i) \cong \chi(-iy) \otimes \beta_4^r(-i)$
for $y, z \in \mathbb{C}^*$.

Apart from the exceptions given above, different pairs of parameters y and z give non-similar representations.

The exceptional cases $n = 5, 6$ is explained by the theory of Hecke-algebra, and will not be dealt with here, as that would require either to much additional theory or the rather ad hoc and lengthy treatment in [7]. These special cases complicate the proof of Theorem 4.5, and it will not be given here.

For $n = 4$, there are two families of additional irreducible representations, one arising via Hecke-algebras as mentioned above, the other from the special relation between B_4 and B_3. We discuss briefly the latter. Denote with π the homomorphism $B_4 \to B_3$ mapping $\sigma_1 \mapsto \sigma_1$, $\sigma_2 \mapsto \sigma_2$ and $\sigma_3 \mapsto \sigma_1$. Then for any representation

τ of B_3 corresponds a representation $\tau\pi$ of B_4. This is a representation because the homomorphism π assures that the braid relations are satisfied. Note that for $n \geq 4$ however, it is not possible to extend a representation of B_{n-1} non-trivially to B_n in this manner, since the braid relation requires generators separated by two or more strings to commute. Thus for $n = 4$ only, there exists the following family of irreducible representations:

$$\tau\pi\colon B_4 \to GL_r(\mathbb{C}) \text{ with } r = 2,\ 3$$

where τ is an irreducible representation of B_3 in $GL_r(\mathbb{C})$ and π is the exceptional homomorphism from B_4 to B_3.

To complete the characterisation above, the irreducible representations $\tau\colon B_3 \to GL_3(\mathbb{C})$ need to be characterised.

Theorem 4.6. *$B_3 = \langle \alpha, \Omega \rangle$ with α, Ω defined as in Section 3.5. Then all irreducible representations $B_3 \to GL_3(\mathbb{C})$ are equivalent to a representation τ of the form below, for some parameters $y \in \mathbb{C}^*$, $a_1, a_2, a_3 \in \mathbb{C}$.*

$$\tau(\alpha) = y^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix}, \quad \tau(\Omega) = y^3(I - AB^T)$$

where $\omega = \exp(\frac{2\pi}{3}i)$ and $A = (a_1, a_2, a_3)^T$, $B = (1,1,1)^T$ with $a_1a_2a_3 \neq 0$ and $a_1 + a_2 + a_3 = 2$.

Proof. Let $\tau\colon B_3 \to GL_3(\mathbb{C})$ be an irreducible representation. Then since both α and Ω lie in the center of B_3, $\alpha^3 = \Omega^2$ and τ is irreducible, we have $\tau(\alpha)^3 = \tau(\Omega)^2 = dI$ for some $d \in \mathbb{C}^*$. Setting $a = \sqrt[3]{d}$, $b = \sqrt{d}$ and $\omega = \exp(\frac{2\pi}{3}i)$, $\tau(\alpha)$ and $\tau(\Omega)$ are then seen to be diagonalisable as follows

$$(\tau(\alpha) - aI)(\tau(\alpha) - a\omega I)(\tau(\alpha) - a\omega^2 I) =$$
$$\tau(\alpha)^3 - a(\omega^2 + \omega + 1)\tau(\alpha)^2 + a^2(\omega^2 + \omega + 1)\tau(\alpha) - a^3 I = \tau(\alpha)^3 - dI = 0$$

since $(\omega^2 + \omega + 1) = 0$, and

$$(\tau(\Omega) - b)(\tau(\Omega) + b) = \tau(\Omega)^2 - dI = 0.$$

Hence the minimal polynomials of $\tau(\alpha)$ and $\tau(\Omega)$ must split into distinct linear factors, and $\tau(\alpha)$ and $\tau(\Omega)$ are diagonalisable. Their possible eigenvalues are $\{a, a\omega, a\omega^2\}$ and $\{b, -b\}$ respectively. Suppose $\tau(\Omega)$ has only one eigenvalue. It is then conjugate to a scalar matrix, in which case τ is not irreducible. Hence $\tau(\Omega)$ has eigenvalues b and $-b$. Suppose $\tau(\alpha)$ has a multiple eigenvalue. Then the corresponding eigenspace is two-dimensional and intersects the two-dimensional eigenspace of $\tau(\Omega)$, implying that $\tau(\alpha)$ and $\tau(\Omega)$ has a common eigenvector and thus that τ is not irreducible. Therefore $\tau(\alpha)$ must have eigenvalues a, $a\omega$, and $a\omega^2$.

As $a^3 = b^2$, set $y = a^{-1}b$ so that $a = y^2$, $b = y^3$, and conjugate such that

$$\tau(\alpha) = y^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix}.$$

Without loss of generality, $\tau(\Omega)$ is conjugate to

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

which is a pseudoreflection. Hence $\tau(\Omega)$ is a pseudoreflection, and we can write $\tau(\Omega) = y^3(I - ab^T)$ for $a = (a_1, a_2, a_3)^T) \in \mathbb{C}^3$ and $b = (b_1, b_2, b_3)^T \in \mathbb{C}^3$ with $b^T a = 2$ (the determinant of $(I - ab^T)$ is $1 - b^T a$). The requirement for τ to be irreducible means that $a_1 a_2 a_3 \neq 0$ and $b_1 b_2 b_3 \neq 0$. It is then possible to conjugate such that $b_1 = b_2 = b_3 = 1$, leaving $\tau(\alpha)$ unaffected. Now, $b^T a = 2$ then implies that $a_1 + a_2 + a_3 = 2$. This is the characterisation in the theorem statement. □

4.6 Unitarisability

Definition 4.16. *A representation $\rho\colon G \to GL_n(\mathbb{C})$ of a group G is called unitary if $\rho(g)$ is unitary for every group elements $g \in G$.*

Unitary representations respect the standard inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{C}^n$, i.e. for a unitary representation ρ, $\langle \rho(g)u, \rho(g)v \rangle = \langle u, v \rangle$ for every group element g and all vectors u and v. As such, unitary representations are those representations preserving the physical observables in quantum mechanics, and thus are the relevant representations when modelling the physics of particle exchange. It is therefore of interest whether a given representation is equivalent to a unitary representation. This property is called unitarisability.

Definition 4.17. *A representation $\pi\colon G \to GL_n(\mathbb{C})$ is called unitarisable if it is similar to some unitary representation.*

Recall that every inner product $(\cdot, \cdot)$ on $\mathbb{C}^n$ may be realized as $\langle \cdot, \cdot \rangle_H$ for some positive definite hermitian matrix H, where $\langle u, v \rangle_H = \langle u, Hv \rangle$. An inner product $\langle \cdot, \cdot \rangle_H$ is said to be invariant under the representation π if $\langle \pi(g)u, \pi(g)v \rangle_H = \langle u, v \rangle_H$ for every group element g and all vectors u and v, i.e. it is invariant under the action of G. It follows from the definition that $\langle \cdot, \cdot \rangle_H$ is invariant under π if and only if $\pi^*(g)H\pi(g) = H$ for every group element g. The existence of an invariant inner product could also be taken as the definition of unitarisability, as the two properties are in fact equivalent.

Lemma 4.3. *A representation $\pi\colon G \to GL_n(\mathbb{C})$ is unitarisable if and only if there is an inner product invariant under π.*

Proof. Assume that such an invariant inner product exists and let H be the associated positive definite hermitian matrix. As such, H may be written in the form $H = P^*P$ with P an invertible matrix. Note that $\pi^* H \pi = H$ implies $\pi H^{-1} \pi^* = H^{-1}$, where the arbitrary group element g is omitted from the notation. Since H is invertible, so is P. Setting $\tilde{\pi} = P\pi P^{-1}$ we have
$\tilde{\pi}^*\tilde{\pi} = (P\pi P^{-1})^* P\pi P^{-1} = (P^*)^{-1}\pi^*(P^*P)\pi P^{-1} = (P^*)^{-1}(\pi^* H\pi)P^{-1} = (P^*)^{-1}HP^{-1} = (P^*)^{-1}P^*PP^{-1} = I$
and
$\tilde{\pi}\tilde{\pi}^* = P\pi P^{-1}(P\pi P^{-1})^* = P\pi P^{-1}(P^*)^{-1}\pi^*P^* = P\pi(P^*P)^{-1}\pi^*P^* = P(\pi H^{-1}\pi^*)P^* = PH^{-1}P^* = I$.
Hence, π is similar to the unitary representation $\tilde{\pi}$, and thus unitarisable.

For the other direction, let $P\pi P^{-1}$ be unitary for some invertible matrix P and set $H = P^*P$. Then $(P\pi P^{-1})^*(P\pi P^{-1}) = I \implies \pi^*P^*P\pi = P^*P$ i.e. $\pi^*H\pi = H$. Hence π preserves the inner product $\langle \cdot, \cdot \rangle_H$. □

Proposition 4.7. *All unitary representations are completely reducible.*

Proof. For a reducible unitary representation ρ of a group G in V, let $A \subset V$ be a proper invariant subspace. Let $B \subset V$ be the orthogonal complement of A in V. Take $a \in A$ and $b \in B$. For any $g \in G$, $0 = \langle a, b\rangle = \langle \rho(g)a, \rho(g)b\rangle$, using that ρ is unitary and A invariant. Hence $\rho(g)b \in B$, i.e. B is invariant and ρ is completely reducible. □

This result holds also for unitarisable representations, since all that was needed was the existence of an invariant inner product, which by Theorem 4.3 is exactly unitarisability. This is stated as a corollary.

Corollary 4.2. *All unitarisable representations are completely reducible.*

For the next theorem, we consider the transposed reduced Burau representation β^T defined by sending the generators σ_i to $\beta(\sigma_i)^T$. Note that for example $\sigma_1\sigma_2$ is sent to $\beta(\sigma_1)^T\beta(\sigma_2)^T$. This is just a translation to a different convention, and the reduced Burau representation could equally well have been defined instead as β^T.

Theorem 4.7. *There exists a hermitian matrix invariant under a rescaling of β^T as defined above.*

Proof. This is due to Squier in [23]. For $|z| = 1$, the hermitian form is

$$H = \begin{pmatrix} \sqrt{z} + \frac{1}{\sqrt{z}} & -1 & 0 & \cdots & 0 \\ -1 & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & -1 & \sqrt{z} + \frac{1}{\sqrt{z}} & -1 \\ 0 & \cdots & 0 & -1 & \sqrt{z} + \frac{1}{\sqrt{z}} \end{pmatrix}$$

and the rescaling is given by

$$P = \begin{pmatrix} \frac{1}{\sqrt{z}} & 0 & & \cdots & 0 \\ 0 & \frac{1}{z} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & 0 & \\ & & & \frac{1}{(\sqrt{z})^{n-1}} & 0 \\ 0 & \cdots & 0 & 0 & \frac{1}{(\sqrt{z})^{n}} \end{pmatrix}.$$

It is easily checked that $P\beta^T P^{-1}$ preserves the hermitian form given by H above. ☐

Letting H_n denote the $(n-1) \times (n-1)$ hermitian matrix associated to β_n, the determinant of H_n may be calculated inductively as $\det(H) = (z^{n-1} + ... + z + 1)/\sqrt{z}^{n-1}$. The roots of this expression are precisely the non-trivial nth roots of unity. Hence we have the following result.

Proposition 4.8. *With $z = \exp(\lambda i)$, the matrix H_n is positive definite for $|\lambda| < \frac{2\pi}{n}$.*

Corollary 4.3. *For z such that H is positive definite, β^T is unitarisable.*

Proof. By Theorem 4.3, it suffices to show that $\forall n \in \mathbb{N}$ there is an inner product invariant under the action of B_n on $\mathbb{C}^n$ given by β^T. By Theorem 4.7, the hermitian matrix H_n defines such an inner product. ☐

Note that although the results concerning unitarisability involve β^T and not the reduced Burau representation as defined in Definition 4.13, the two are similar by Theorem 4.5. Thus the reduced Burau representation is unitarisable by the above.

The form of the fully reduced Burau representation suggests that unitarisability of β should imply unitarisability of β^r. More generally, restricting the invariant inner product to a subrepresentation of any unitarisable representation should yield a unitarisation of the subrepresentation.

Proposition 4.9. *A subrepresentation of a unitarisable representation is unitarisable.*

Proof. Let π be a unitarisable representation V, H the associated hermitian matrix and $\hat{\pi}$ a subrepresentation on the invariant subspace $U \subset V$. By Corollary 4.2, π is completely reducible, and may be written on the form

$$\pi = \hat{\pi} \oplus R$$

for some subrepresentation R of π on the complement of U. Now $\beta^* H \beta = H$ imply

$$\hat{\pi}^* H_1 \hat{\pi} = H_1 \quad \text{and} \quad R^* H_2 R = H_2$$

where H_1 and H_2 are the restrictions of H to U and the complement of U respectively. As H is a positive definite hermitian matrix, so is H_1. Hence H_1 defines an invariant inner product on U and $\hat{\pi}$ is unitarisable. □

Corollary 4.4. *The fully reduced Burau representation is unitarisable whenever the reduced Burau representation is unitarisable.*

Theorem 4.5 imply that for $n \geq 7$, completely reducible representations of B_n of degree greater than 1 and less than $n-2$ are similar to representations sending all generators to the same diagonal matrix. Such representations are called diagonal in the proposition below. This follows since unitarisability implies complete reducibility of the Burau type representations, and for degrees less than $n-2$ the irreducible components must be one-dimensional. If the representation is completely reducible, it thus reduces to a direct sum of one-dimensional representations. By the examination of such representations in Section 4.2, all generators map to the same element. This applies specifically in the case of anyons, where the representation are required to be unitary and hence completely reducible.

As a final example, the transposed reduced Burau representation of degree two of B_3 is unitarised using Lemma 4.3, thus obtaining an expression for the irreducible representations of B_3 in $U(2)$.

The eigenvalues and orthonormal eigenvectors of H from Theorem 4.7 are

$$\begin{aligned} \lambda_1 &= \sqrt{z} + \tfrac{1}{\sqrt{z}} - 1, \quad v_1 = \tfrac{1}{\sqrt{2}}(1,1)^T \\ \lambda_2 &= \sqrt{z} + \tfrac{1}{\sqrt{z}} + 1, \quad v_2 = \tfrac{1}{\sqrt{2}}(1,-1) \end{aligned}$$

and H may be diagonalised as

$$H = \begin{pmatrix} \sqrt{z} + \frac{1}{\sqrt{z}} & -1 \\ -1 & \sqrt{z} + \frac{1}{\sqrt{z}} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \sqrt{z} + \frac{1}{\sqrt{z}} - 1 & 0 \\ 0 & \sqrt{z} + \frac{1}{\sqrt{z}} + 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} .$$

Setting

$$P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \sqrt{\sqrt{z}+\frac{1}{\sqrt{z}}-1} & 0 \\ 0 & \sqrt{\sqrt{z}+\frac{1}{\sqrt{z}}+1} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{\sqrt{z}+\frac{1}{\sqrt{z}}-1} & \sqrt{\sqrt{z}+\frac{1}{\sqrt{z}}-1} \\ \sqrt{\sqrt{z}+\frac{1}{\sqrt{z}}+1} & -\sqrt{\sqrt{z}+\frac{1}{\sqrt{z}}+1} \end{pmatrix}$$

we obtain $H = P^*P$.

After rescaling as in Theorem 4.7, β^T is given on the two generators of B_3 as

$$\begin{aligned} \beta^T(\sigma_1) &= \begin{pmatrix} -z & \sqrt{z} \\ 0 & 1 \end{pmatrix} \\ \beta^T(\sigma_2) &= \begin{pmatrix} 1 & 0 \\ \sqrt{z} & -z \end{pmatrix} \end{aligned}.$$

With the matrix P as above, setting $\tilde{\beta} = P\beta^T P^{-1}$ yields according to Lemma 4.3 a unitary representation, and indeed

$$\tilde{\beta}(\sigma_1) = \frac{1}{2} \begin{pmatrix} -z+\sqrt{z}+1 & -\sqrt{z}\sqrt{\sqrt{z}+\frac{1}{\sqrt{z}}-1}\sqrt{\sqrt{z}+\frac{1}{\sqrt{z}}+1} \\ -\sqrt{z}\sqrt{\sqrt{z}+\frac{1}{\sqrt{z}}-1}\sqrt{\sqrt{z}+\frac{1}{\sqrt{z}}+1} & -z-\sqrt{z}+1 \end{pmatrix}$$

and

$$\tilde{\beta}(\sigma_2) = \frac{1}{2} \begin{pmatrix} -z+\sqrt{z}+1 & \sqrt{z}\sqrt{\sqrt{z}+\frac{1}{\sqrt{z}}-1}\sqrt{\sqrt{z}+\frac{1}{\sqrt{z}}+1} \\ \sqrt{z}\sqrt{\sqrt{z}+\frac{1}{\sqrt{z}}-1}\sqrt{\sqrt{z}+\frac{1}{\sqrt{z}}+1} & -z-\sqrt{z}+1 \end{pmatrix}.$$

are unitary.

Chapter 5

Summary

Non-abelian anyons correspond to non-trivial unitary representations of the braid group of degree two and higher. These have been examined in the case where the degree of the representations is limited above by the number of particles, and the unitarisability of such representations has been investigated. In Sections 4.3 and 4.5, specific cases of representations of degree two and three respectively of B_3 was characterised.

In Sections 4.5, the main conclusion was the following: For a fixed number of particles n, with $n > 6$, the irreducible representations of the corresponding braid group are either one-dimensional or of degree greater than or equal to $n-2$. In the case that the representation is of degree $n-2$ or $n-1$, it is of Burau type, as expressed in Theorem 4.5. Additionally, as shown in Section 4.6, representations of Burau type are unitarisable for certain specialisations of the parameter z, and thus relevant for anyon statistics. Furthermore, for $n > 6$, the only possible unitary representations of degree $2 \leq k < n-2$ are similar to diagonal representations, and hence abelian. In other words, for more than 6 particles, non-abelian anyons correspond to representations, and hence wave functions, of dimension close to or greater than the number of particles.

Bibliography

[1] M.A. Armstrong, *Basic Topology*, Springer, 1983.

[2] E. Artin, *Theorie der Zöpfe*, Hamburg Abh. 4 (1925), 47-72.

[3] E.Artin, *Theory of braids*, Ann. of Math. 48 (2) (1947) 101-126.

[4] J.E. Avron, O. Kenneth, *Braiding fluxes in Pauli Hamiltonian*, Ann. of Phys. 349 (2014) 325-349.

[5] W. Burau, *Über Zopfgruppen und gleichsinnig verdrillte Verkettungen*, Abh. Math. Sem. Univ. Hamburg 11 (1936) 179-186.

[6] D. Dummit, R. Foote, *Abstract Algebra* 3rd ed., Wiley, 2003.

[7] E. Formanek, *Braid group representations of low degree*, Proc. London Math. Soc. (3) 73 (1996) 279-322.

[8] A. Hatcher, *Algebraic Topology*, Cambridge University Press, 2002.

[9] C. Kassel, V. Turaev, *Braid Groups*, Springer, 2008.

[10] L.H. Kauffman, S.J. Lomonaco Jr., *Topological Quantum Computing and SU(2) Braid Group Representations*, Proc. of SPIE Vol. 6976 69760M-2 (2008).

[11] A. Kitaev, *Topological Quantum Codes and Anyons*, Proc. of Symp. in Applied Math. (2000).

[12] R. B. Laughlin, *Nobel Lecture: Fractional quantization*, Rev. Mod. Phys. 71 (1999) 863-874.

[13] J.M Leinaas, J. Myrheim, *On the Theory of Identical Particles*, Il Nuovo Cimento Vol. 37 B, N.1 (1977).

[14] V. Lomonosov, P. Rosenthal, *The simplest proof of Burnside's theorem on matrix algebras*, Linear Algebra and its Applications 383 (2004) 45-47.

[15] D. Lundholm, J. P. Solovej, *Hardy and Lieb-Thirring inequalities for anyons*, Commun. Math. Phys. 322 (2013) 883-908.

[16] D. Lundholm, J. P. Solovej, *Local exclusion principle for identical particles obeying intermediate and fractional statistics*, Phys. Rev. A 88 (2013) 062106.

[17] D. Lundholm, J. P. Solovej, *Local exclusion and Lieb-Thirring inequalities for intermediate and fractional statistics*, Ann. Henri Poincaré 15 (2014) 1061-1107.

[18] J. Myrheim, *Anyons*; in Topological aspects of low dimensional systems (Les Houches, 1998), pp. 265–413, EDP Sci., Les Ulis, 1999.

[19] M. Nakahara, *Geometry, topology and physics*, Institute of Physics Publishing, 2003.

[20] C. Pinter, *A Book of Abstract Algebra* 2nd Ed., Dover, 2009.

[21] J. Preskill, *Quantum Computation*, Lecture Notes for Physics 219, California Institute of Technology, 2004.

[22] S. Rao, *An Anyon Primer*, http://arxiv.org/abs/hep-th/9209066v3, 2001.

[23] C. Squier, *The Burau representation is unitary*, Proc. Amer. Math. Soc. 90 (1984) 199-202.

[24] Z. Wang, *Topological Quantum Computation*, Regional Conference Series in Mathematics / Conference Board of the Mathematical Sciences No. 112, 2010.

[25] F. Wilczek, *Quantum Mechanics of Fractional-Spin Particles*, Phys. Rev. Lett 49 (1982) 957-959.